Kevin Nitièma | 1002673298
INF2178H: Experimental Design for Data Science – Assignment 4

MRI Study: Demographic Influence on Dementia Brain Health

1. Introduction

Understanding the factors influencing brain health and cognitive function is crucial in the context of aging and neurodegenerative diseases such as dementia. In this report, we examine a longitudinal study on MRI results of patients with and without dementia to explore the relationships between diagnosis, gender disparities, and brain volume changes over time. Our analysis aims to address two fundamental research questions:

- **[R1]:** How does diagnosis status (Demented vs Nondemented) influence the brain over time?
- **[R2]:** How does gender influence changes in demented patients' brain volume over time?

2. Exploratory Data Analysis (EDA)

After cleaning our dataset, we started by exploring its numerical aspects and representing them visually through various plots. Here, we used box plots, illustrated in *Figure 1* and *Figure 2*, to assess the distributions of `Age` and `EDUC` (years of education) among the study subjects. These visualizations revealed a **normal distribution among older adults**, and an average of **14.5 years of education** among subjects.

Next, upon examining the correlation matrix heatmap *(Figure 3)*, we observed a **significant negative correlation** between `SES` and `EDUC` $(r = -0.73)$, `MMSE` and `CDR` $(r > -0.7)$, and finally `ASF` and `eTIV` $(r > -0.99)$. Assessing these relationships later contributed to more meaningful feature selection (e.g. `ASF`) in our analysis.

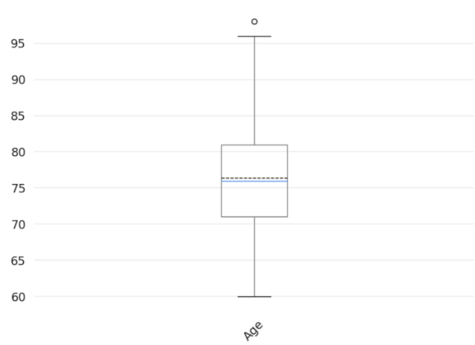

Figure 1: Box plot of Subjects' Age

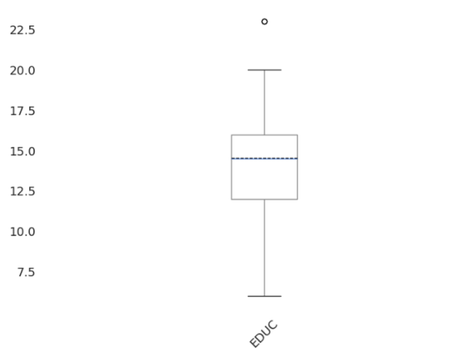

Figure 2: Box plot of Subjects' Years of Education

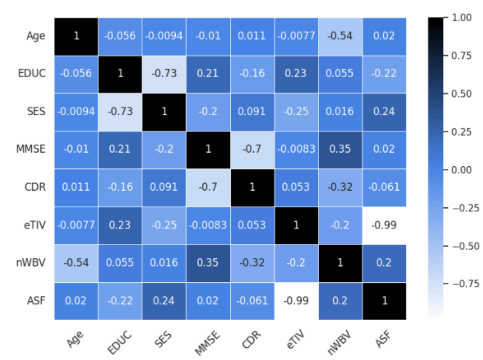

Figure 3: Correlation Matrix Heatmap

Next, we chose to examine the distribution of variables that could serve as potential dependent variables for our study. To do so, we plotted histograms *(Figure 4)* for **Normalize Whole Brain Volume** (`nWBV`), **Atlas Scaling Factor** (`ASF`), and **Estimated Total Intracranial Volume** (`eTIV`) These measures are frequently used in neuroimaging studies, especially in the analysis of MRI data, as they represent normalized/standardized indicators of brain size relative to a template. We noticed a **consistent normal distribution** among most continuous variables, which we consider as ideal candidates for our future analysis.

We also noted significant left skewness in the distribution of **Mini Mental State Examination** (`MMSE`), which serves as a screening tool for cognitive impairment/dementia. Despite its potential relevance to our study, we chose to exclude it due to its **departure from normality** *(Figure 4 - far right)* and the constraints posed by our **limited sample size.**

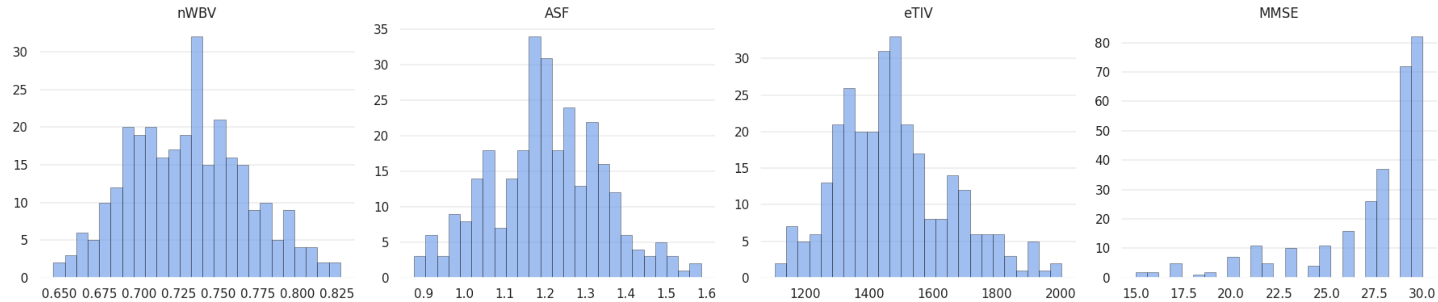

Figure 4: Frequency distribution of variables relevant to neuroimaging studies

After thoroughly examining our dataset, our discoveries led us to two (2) critical research questions. These questions will help deepen our understanding of how (1) **diagnosis status** and (2) **gender** influence brain health over time.

3. Brain Volume Variation in Dementia: Impact of Diagnosis

[R1]: How does diagnosis status (Demented vs Nondemented) influence the brain over time?

To explore the contribution of different levels of diagnosis status (`Group`) to the variation in brain volume (`nWBV`) between the subject's first MRI scan during `Visit` 1 and their second scan a year later at `Visit` 2 (see *Figure 7*), a **Mixed-ANOVA** model was conducted with the following variables:

- **Dependent Variable:** `nWBV` (brain volume)
- **Between-subjects:** `Group` (diagnosis status: *`Demented`* or *`Nondemented`*)
- **Within-subjects:** `Visit` (*`Visit 1`* and *`Visit 2`*)

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
Group	0.029	1	129	0.029	12.263	<0.001	0.087	NaN
Visit	0.006	1	129	0.006	80.305	<0.001	0.384	1
Interaction	0.000	1	129	0.000	2.538	0.114	0.019	NaN

Figure 5: Mixed-ANOVA Results (Study 1)

Mixed-ANOVA Results:
Our analysis *(Figure 5)* revealed that both the Group and Visit variables on brain volume are statistically significant with both $p\_values < 0.001$, with $\alpha = 0.05$. This indicates that there are **significant differences in brain volume** between individuals diagnosed as Demented versus Nondemented and that there are significant changes in brain volume over time, from Visit 1 to Visit 2, **regardless of diagnosis status**. However, there is no significant interaction effect ($p\_value = 0.114$), indicating that the relationship between diagnosis status and changes in brain volume over time is **similar across both Demented and Nondemented** individuals.

Post Hoc Results:
Next, we conducted a post hoc *(Figure 6)* study which confirmed the significant main effects of `Visit` and `Group` as well as the lack of significant interaction between `Visit` and `Group` (all $p\_values = 0.001$). Hence, both groups (Demented and Nondemented) showed major changes in brain volume over time, but the **differences between them remained consistent**, as displayed in *Figure 8*.

Contrast	Visit	A	B	Paired	Parametric	T	dof	alternative	p-unc	p-corr	BF10
Visit	-	1	2	TRUE	TRUE	8.909	130.000	two-sided	<0.001	NaN	1.62E+12
Group	-	Demented	Nondemented	FALSE	TRUE	-3.531	128.734	two-sided	0.001	NaN	45.890
Visit * Group	1	Demented	Nondemented	FALSE	TRUE	-3.248	128.422	two-sided	0.001	0.001	20.163
Visit * Group	2	Demented	Nondemented	FALSE	TRUE	-3.702	129.000	two-sided	<0.001	0.001	77.768

Figure 6: Post-Hoc Test Results (Study 1)

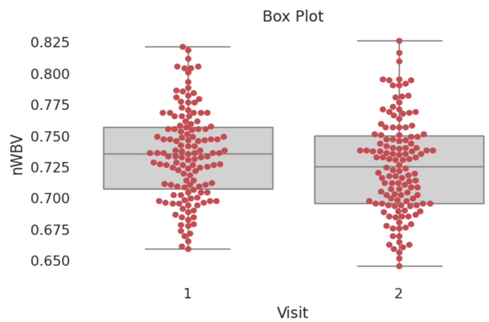

Figure 7: Box Plot of nWBV for Visit 1 and Visit 2

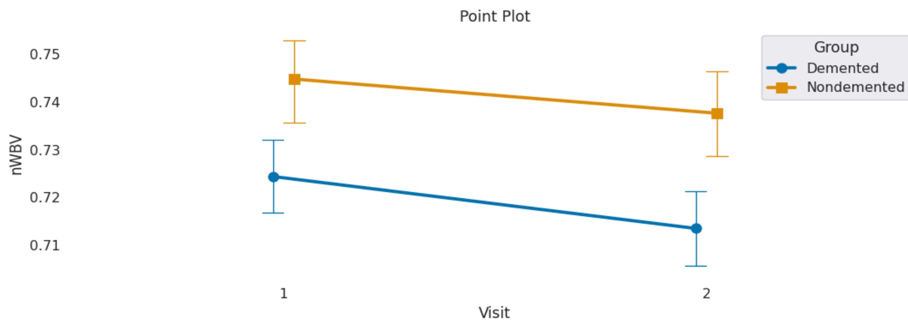

Figure 8: Interaction Plot - Demented vs.Nondemented across Visits

Confirming Assumptions:

- ✔ **Sphericity:** Mauchly's Test of Sphericity yielded $p\_value \approx 1$ with $\alpha = 0.05$.
- ✔ **Normality:** Shapiro-Wilk tests yielded $p\_values > 0.2$ with $\alpha = 0.05$. (also confirmed by *Figure 7*)

Now, we transition from examining differences between Demented and Nondemented subjects to investigating demographic factors, such as gender, and their impact on brain health in demented patients over time.

4. Gender Disparities in Brain Volume Trajectories in Dementia

[R2]: How does gender influence changes in demented patients' brain volume over time?

To address this question, we conducted another **Mixed-ANOVA study** which followed a similar format as our previous study 1. Our model's variables are as follows:

- **Dependent Variable:** `ASF` (brain volume)
- **Between-subjects:** `Gender` (*`M`* or *`F`*)
- **Within-subjects:** `Visit` (*`Visit 1`* and *`Visit 2`*)

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
Gender	0.711	1.000	60.000	0.711	26.487	<0.001	0.306	NaN
Visit	0.003	1.000	60.000	0.003	5.301	0.025	0.081	1
Interaction	0.000	1.000	60.000	0.000	0.177	0.675	0.003	NaN

Figure 9: Mixed-ANOVA Results (Study 2)

Mixed-ANOVA Results:
Here, our results suggest that there are **significant differences in brain volume between male and female demented patients** ($p\_value < 0.001$), with females tending to have higher brain volumes. Additionally, there are significant changes in brain volume over time within demented patients, but there is **no significant interaction between gender and changes in brain volume over time** ($p\_value = 0.075$). Initially, it seemed that gender had a consistent effect on brain volume changes over time in demented patients, also indicated in *Figure 11*. However, our **post hoc test** revealed otherwise, demonstrating the presence of interactions that were not initially apparent.

Post Hoc Results:
Our post hoc test confirms the significant main effects of `Visit` and `Gender`, as well as the significant interaction effect between `Visit` and `Gender`. Most importantly, the significant interaction effect between `Visit` and `Gender` indicates that the relationship between changes in brain volume over time and gender is **not consistent across different time points**. This contradicts initial findings and suggests (with more reliability) that **gender influences the trajectory of changes in brain volume over time** within the demented population.

Contrast	Visit	A	B	Paired	Parametric	T	dof	alternative	p-unc	p-corr	BF10
Visit	-	1	2	TRUE	TRUE	2.318	61.000	two-sided	0.024	NaN	1.650
Gender	-	F	M	FALSE	TRUE	5.140	57.428	two-sided	<0.001	NaN	4724.372
Visit * Gender	1	F	M	FALSE	TRUE	5.203	57.581	two-sided	<0.001	5.5E-06	5827.682
Visit * Gender	2	F	M	FALSE	TRUE	4.962	57.136	two-sided	<0.001	6.6E-06	2631.553

Figure 10: Post-Hoc Test Results (Study 2)

Confirming Assumptions:

- ✔ **Sphericity:** Mauchly's Test of Sphericity yielded $p\_value \approx 1$ with $\alpha = 0.05$.
- ✔* **Normality:** Shapiro-Wilk tests yielded $p\_values = 0.1$ with $\alpha = 0.05$, except for the female group, likely due to its small sample size. Increasing the sample size will likely restore normality.

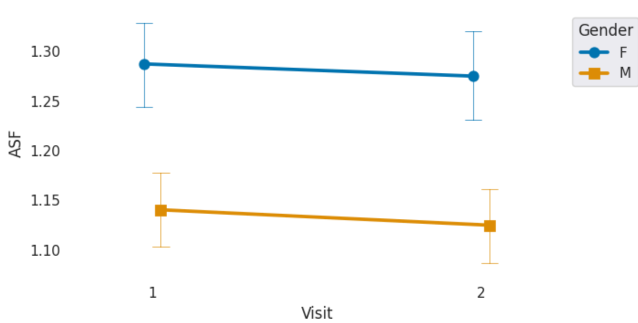

Figure 11: Interaction Plot - Female vs.Male across Visits

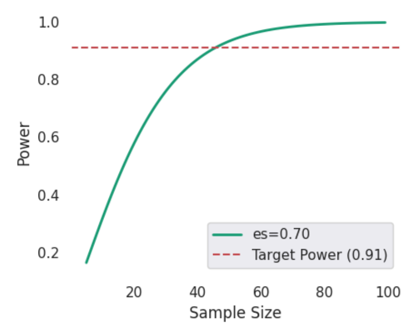

Figure 12: Statistical Power Analysis for t-Tests

5. Statistical Power Analysis

In our theoretical experiment (*Figure 12*), we conducted a **statistical power analysis for t-tests** to determine the appropriate sample size. With a desired power of 0.91, alpha level of 0.05, and effect size of 0.7, our analysis revealed that a sample size of 46 would provide sufficient statistical power to detect the specified effect.

6. Conclusion

Our comprehensive study into the influences on brain volume changes over time have yielded insightful results. First, in Study 1, we observed **significant variations in brain volume between individuals diagnosed as Demented vs Nondemented**, as well as major changes in brain volume over time for both groups. Study 2, revealed significant differences in brain volume between **male and female demented patients**, with females exhibiting higher brain volumes. Additionally, the interaction effect between gender and changes in brain volume over time suggested that **gender influences the trajectory of brain volume changes within the demented population**. Nonetheless, further study should be conducted with **larger datasets** to enhance our understanding of demographics and their impact on neurodegenerative diseases such as Dementia.